企业级Hadoop平台源码二次开发-第三天

一、课前准备

1.掌握上一次课内容

二、课堂主题

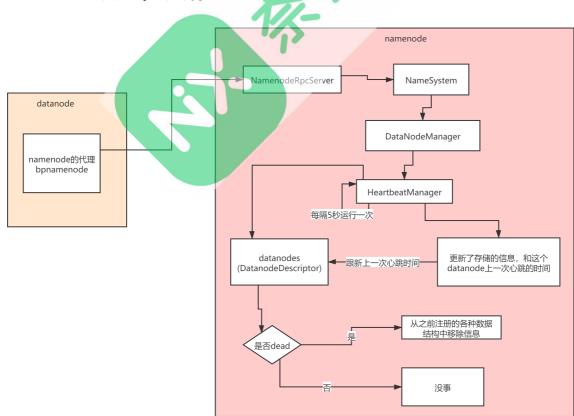
剖析NameNode管理元数据原理,学习优秀开源技术的架构设计思想,同时为后面进行二次开发打下 基础

三、课程目标

- 1. HDFS目录树管理
- 2. HDFS元数据双缓存方案深度剖析
- 3. HDFS元数据写JournalNode流程分析
- 4. StandbyNameNode Checkpoint原理深度剖析
- 5. ActiveNameNode FSimage文件更新流程分析

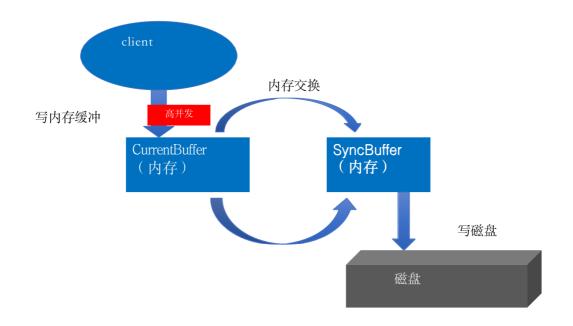
四、知识要点



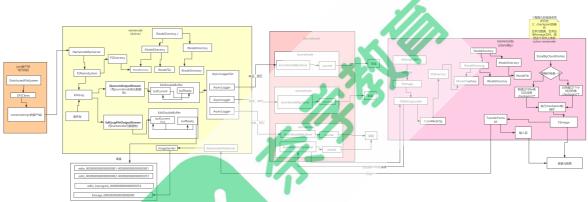


4.2 HDFS元数据管理流程 (150分钟)

4.2.1 HDFS双缓冲机制实现 (30分钟)



4.2.2 HDFS元数据管理流程源码剖析 (120分钟)



五、总结(5分钟)

- 1. HDFS目录树管理
- 2. HDFS元数据双缓存方案深度剖析
- 3. HDFS元数据写JournalNode流程分析
- 4. StandbyNameNode Checkpoint原理深度剖析
- 5. ActiveNameNode FSimage文件更新流程分析

六、作业

- 1. 画一遍HDFS生成元数据的流程图
- 2. 画一遍HDFS写数据的流程图

七、互动